3D OBJECT MODEL FITTING TO STILL IMAGES USING LINEAR COMBINATION METHOD OF 2D ASPECT IMAGES

Hiroyasu Sakamoto and Masahide Kawakami
Kyushu Univ. of Design — Namko Co.
4-9-1, Shiobaru, Minami-ku, Fukuoka, 815 JAPAN.
1-1-32, Shin-urashima, Kanagawa-ku, Yokohama, 221 JAPAN

ABSTRACT

This paper describes a method for fitting 3D object model to still (single) 2D observed image by searching for the model's optimum posture parameters in the parameter space through LMM iteration method (a stabilized Gauss-Newton method). In the method, we propose a new approach to feeding an excellent initial parameter value to the iteration method via aspect identification by applying linear combination method of 2D aspect imaged[1], and via restricting the feature correspondence between the 2D image and the model's projection image to their circumferential features. Due to these schemes, the method gives a simple, fast and robust procedure. Numerical simulation using 500 randomly postured synthetic images of an irregular pentagonal prism shows 100% correct match only in 2.98 average iteration counts of the LMM method.

1. Introduction

It is one of the central tasks in 3D model based computer vision to find out posture parameters (rotational angles and translational position) of an object by fitting 3D object model's projection image (3D image) to bserved 2D gray scale image (2D image). The fitting problem has been studied mainly in two different basic strategies; qualitative or topological fitting [2],[3] and quantitative or spatial fitting [4],[5]. Compared to the former, it is reported that the latter yields smaller number of multiple solutions (typically unique solution if the object has no symmetry) and tends to be robuster in the case where occlusion and noise exist in the observations. The spatial fitting is commonly formulated in a nonlinear minimization problem, and an iterative method such as Gauss-Newton (GN) method is frequently employed on searching for the optimum parameters in the parameter space.

Lowe [4] employed GN method in spatial model fitting problem successfully. Due to the existence of multiple local minima of error evaluating function, his method has to run GN iteration many times using different starting (initial) parameters. Even after many runs which consume a long time, it is possible that global minimum can not be found. Little is described how to choose initial parameters.

This method is extended to image frame sequence for motion tracking [5]. Levenberg-Marquardt (LM) method supersedes GN method with a few more improvements. The result of the previous frame can successfully be adopted as initial parameters for intermediate frames, but there is no improvement for estimating initial parameters of the starting frame.

One of main problems of the above and other conventional 3D model fitting methods seems to be absence of or difficulties in finding for such a good initial parameter value estimation that leads the iteration procedure to the global minimum. This is a consequence of unknown feature correspondence between 2D images and the model, and of a large number of the possible correspondences.

This paper proposes a new method of quantitatively determining posture parameters of an object from its single 2D image. The method gives an estimation scheme of good initial parameters by analyzing aspect of the 2D image and by making correspondence between only circumferential features of the 2D image with those of the 3D image. We introduce linear combination method of 2D images by Ullman and Basri [1] for aspect identification. In order to stabilize the solution, to achieve faster converging speed and to make the method robust to noisy and defective 2D image, we employ Levenberg-Marquardt-Morrison (LMM) iteration method with two kinds of error measure between the two images; one is sum of squared vertex Euclidean distances and the other is sum of squared edge distances.

In section 2, the initial parameter estimation scheme is described with the 2D image linear combination method and a rule for making correspondence between features of the two images. The LMM iteration method and two kinds of error measure are explained in section 3. Results of a numerical simulation are shown in section 4.

2. Initial parameter estimation

In all iterative methods, it is essentially important to estimate such initial parameter values that can most probably converge to the global minimum. Observed aspect of the 2D image must be an effective cue for the estimation, and a representative view of the aspect must give an excellent initial parameter for our iteration method.

2.1 Linear combination method of 2D image

Ullman and Basri have reported linear combination method of 2D aspect images which are obtained by orthographic projection [1]. On the basis of their method, it is possible to identify the aspect of the observed image using multiplications of aspect identifier matrices with its feature vectors.

Let $\boldsymbol{x}$ be n dimensional x-coordinate vectors of n feature points of an observed 2D image. This is represented as a linear combination of x-coordinate vectors of 2D bases images' features $\boldsymbol{x}_i$ $(i = 1,2,3)$ whose elements are arranged in the same order as $\boldsymbol{x}$ (the arranging method will be described later) and translation basis vector $\boldsymbol{x}_4 = [1,\cdots,1]^T$. Here, ^{T} indicates a transpose. Using the same notations, this holds also for y-coordinates.

$$\boldsymbol{x} = \sum_{i=1}^{4} a_i \boldsymbol{x}_i, \quad \boldsymbol{y} = \sum_{i=1}^{4} b_i \boldsymbol{y}_i . \qquad (1),(2)$$

Adopting these bases vectors, it is possible to construct $n \times n$ aspect identifier matrices $\boldsymbol{L}_X$ (for x-coordinates) and $\boldsymbol{L}_Y$ (y-coordinates) of the aspect. Suppose that $\boldsymbol{L}_X$ satisfies next equation.

$$\boldsymbol{L}_X \boldsymbol{X} = \boldsymbol{X}_0, \qquad (3)$$

where, choosing vectors $\boldsymbol{x}_i$ $(i = 5,\cdots,n)$ by Schmidt's method so that they are orthogonal to $\{\boldsymbol{x}_i \mid i = 1,\cdots,4\}$ and also to each other, and using an arbitrary vector $\boldsymbol{x}_0$, $\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \boldsymbol{x}_4, \boldsymbol{x}_5, \cdots, \boldsymbol{x}_n]$, $\boldsymbol{X}_0 = [\boldsymbol{x}_0, \boldsymbol{x}_0, \boldsymbol{x}_0, \boldsymbol{x}_0, \boldsymbol{x}_5, \cdots, \boldsymbol{x}_n]$. Then, $\boldsymbol{L}_X$ maps aribitrary $\boldsymbol{x}$ of Eq(1) to $\Sigma_i\, a_i\, \boldsymbol{x}_0$, a vector proportional to $\boldsymbol{x}_0$. For the purpose of aspect identification, it is convenient to put $\boldsymbol{x}_0 = \boldsymbol{0}$. The same is true also for y-coordinates. Thus, after preparing $\boldsymbol{L}_X = \boldsymbol{X}_0 \boldsymbol{X}^{-1}$ and $\boldsymbol{L}_Y = \boldsymbol{Y}_0 \boldsymbol{Y}^{-1}$ for each possible aspect, it is easy to check which aspect is observed in the image by evaluating $A = \|\boldsymbol{L}_X \boldsymbol{x}\| + \|\boldsymbol{L}_Y \boldsymbol{y}\| \geq 0$. Because $\boldsymbol{x}_0 = \boldsymbol{0}$, we pick up two aspect candidates for the 2D image which show the smallest values of A.

Using aspect analysis method by Ikeuchi and Kanade [6], we obtain total number of aspects and their representative views as those observed from centers of gravity of the aspect regions on Gaussian sphere. Three bases images of each aspect are selected from the aspect region so that they are as much linear independent as possible, and corresponding aspect identifier matrices $\boldsymbol{L}_X$ and $\boldsymbol{L}_Y$ are calculated.

2.2 Aspect and initial parameter

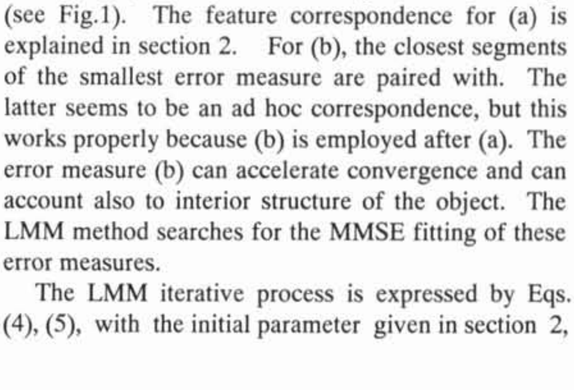
Because of Gaussian sphere, the representative view of each aspect gives the initial values of rotation angles θ_x and θ_y about x and y axes on the image plane. On the other hand, rotation of the object about z (lens) axis gives no aspect change. Thus, initial value of θ_z is related to feature point correspondence between 2D and 3D images within the aspect. In order to drastically simplify this problem at the initial step of iteration, we restrict the correspondence between features to circumferential ones of the 2D and 3D images. In actual image acquisition, it is reasonable to believe that circumferential features are more reliable than interior ones. If we have n such features, there are n correspondences between 2D and 3D images depending upon from which feature we start to pick up. This gives an initial estimate of θ_z.

Initial values of scaling and (x, y) translations of the object are estimated by variance and mean of the 2D object image area. Total parameters number is 6.

3. Error measures and LMM iteration

In order to get faster and stable convergence of LMM method, we employ two kinds of error measures between the two images ; (a)differences of x and y coordinates of circumferential vertices between the two images, and (b) longitudinal and transversal distances to a line segment of 3D image measured from a few points on each line segment of 2D image (see Fig.1). The feature correspondence for (a) is explained in section 2. For (b), the closest segments of the smallest error measure are paired with. The latter seems to be an ad hoc correspondence, but this works properly because (b) is employed after (a). The error measure (b) can accelerate convergence and can account also to interior structure of the object. The LMM method searches for the MMSE fitting of these error measures.

The LMM iterative process is expressed by Eqs. (4), (5), with the initial parameter given in section 2,



Error measure (a), Error measure (b).

Figure 1. Two error measures between line segments of 2D image and 3D image.

where p_k, d_k, e_k and J_k represent 6-dimensional parameter vector, correction vector, m-dimensional error measure vector, and $m\times6$ Jacobian whose (i,j)-th element is $[\partial e_{k,i}/\partial p_{k,j}]$, respectively at k-th iteration step. Identity matrix (6×6) I with adjustable positive scalar λ_k stabilizes the convergence.

$$p_{k+1} = p_k - d_k \ , \tag{4}$$

where d_k is the MS solution of Eq.(5).

$$\begin{bmatrix} J_k \\ \lambda_k I \end{bmatrix} d_k = \begin{bmatrix} e_k \\ 0 \end{bmatrix}. \tag{5}$$

Because imaging of orthogonal projection is obtained by product of translational and rotational matrices, calculation of partial derivative with respect to each parameter reduces to modification of each matrix. Therefore, calculations of J_k for our error measures are simple.

Using the error measure (a), the procedure is iterated until the condition $\|e_{k+1} - e_k\| / m < 10^{-4}$ is satisfied. Then, we switch the error measure to (b), and repeat until $\|e_k\| / m < 10^{-6}$ is satisfied. If both convergence conditions were not satisfied within predetermined iteration counts, the second aspect candidate is adopted.

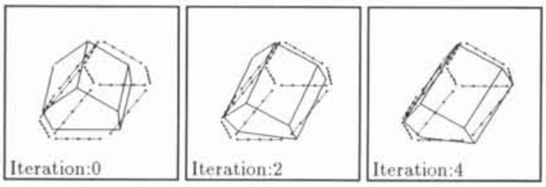
4. Numerical simulation

We employ an irregular pentagonal prism who has 21 aspects in this simulation. Five hundreds of 2D observation images are synthesized by generating posture parameters of uniform random variable. Small observation noise (0.3% of the object size) is introduced in synthesizing all 2D images.

Fig. 2 shows an example of false converging process of the 3D model image (dotted lines) to the 2D image (real lines with circle points on the edges) when the process startes from a view of different aspect using error measure (a). After 6 iterations the procedure converges to a local minima. Aspect of the final image has changed from the starting aspect. Though the procedure can traverse aspect borders, the true aspect is too far from the starting aspect to be reached. This example shows significance of the initial parameter and that its aspect should neighbor the true one.

Fig. 3 shows an example of successful fitting which starts from the same aspect's representative view. Measured error values are written in each windows. In this case, it took only two iterations using error measure (a) exclusively. The convergence condition for error measure (b) is also satisfied finally.

Next example depicted in Fig. 4 is the case where 2D observed image is very close to aspect changeover. In such cases, aspect identification, feature correspondence, error measurement and consequently convergence are all difficult. After 3 iteration (v1~v3) with error measure (a) and an iteration (e1) with error measure (b), error decreasing rate becomes very slow, and there still remains small amount of error. The iteration process will not converge any more. Because of the small noise introduced in 2D observation image, there is mistake in aspect identification. Thus the second aspect candidate is adopted, and after the same iteration count as the first aspect, the 3D model image fits correctly.



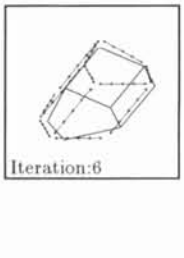


Figure 2. An example of LMM iteration process starting from a false aspect.

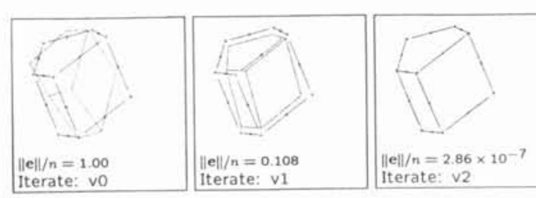


Figure 3. An example of correct fitting after 2 iterations started from the identical aspect.

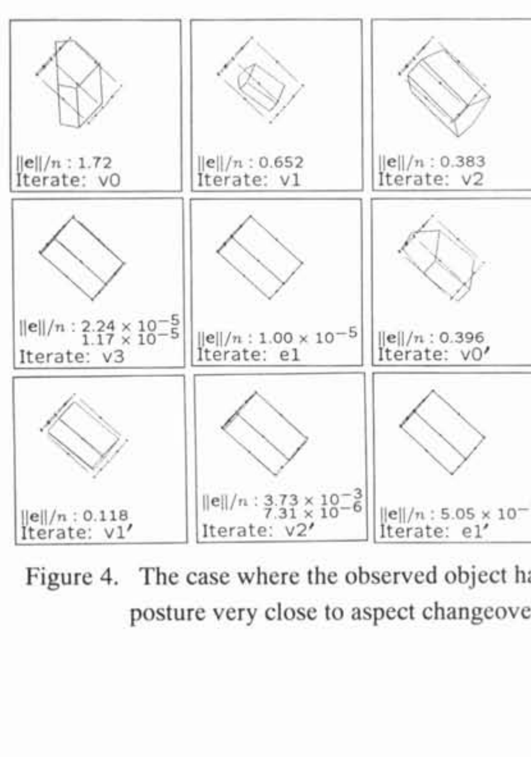


Figure 4. The case where the observed object has posture very close to aspect changeover.

All the 500 2D synthesized images have had 100% correct fitting, among which 490 have succeeded with the first aspect candidate and error measure (a), 6 with the first candidate and error (b), and 4 with the second one and error (b). The histogram of total iteration counts is depicted in Fig. 5. Note that thick bars at every 5 iteration count are grid of the graph. The minimum, maximum, mean and standard deviation of all 500 iteration counts are 1, 42, 2.98 and 3.34, respectively. For the cases where correct fitting is obtained by the first aspect candidate and the second candedate, the average values of aspect identification measure $A = \|\boldsymbol{L}_X\boldsymbol{x}\|+\|\boldsymbol{L}_Y\boldsymbol{y}\|$ of observed 2D images are 2.8e-5 (to the first candidate), and 4.5e-4 (to the first one) and 2.8 (to the second one), respectively.

The above results show the importance of aspect information for the initial posture parameter. If the iteration starts from a far different aspect, it will surely converge to a local minimum. When a wrong initial parameter is applied for an image sequence, it is possible that the method can trace the whole sequence in failure.

On the other hand, initial parameter of identical aspect view can very quickly lead the iteration to the global minimum. Circumferential correspondence of features and two kinds of error measure simplify the method and accelerate the convergence speed.

Difficulty is the 2D image positioned very close to aspect changeover. It needs larger iteration counts and the second aspect candidate. However, it is shown that the second aspect candidate can give very effective initial parameter estimation.

The most time consuming parts of the method are the aspect identification and the calculation of Jacobian. But, current micro computers can calculate and wireframe graphic display of our method in real time.

5. Conclusions

A quantitative method for fitting 3D object model image to single observed 2D image is discussed, and

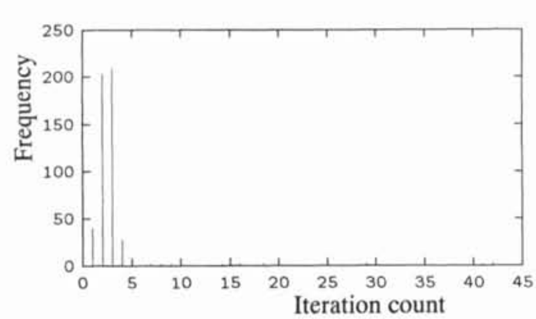


Figure 5. The histogram of 500 iteration counts.

a new method for initial parameter estimation for the 2D image is proposed. The initial parameter estimated by aspect identification and two kinds of error measure make LMM iteration very fast and robust. It is pointed out that when the observed image is close to aspect changeover, the problem becomes very difficult. However, the second aspect candidate gives a very good estimation for such difficult cases.

Although a small amount of noise is included in our simulation data, we need to check the performance of the method to much stronger noise and occlusion. Real images are also necessary for the checking. The method should be extended to objects with self-occluding contours.

REFERENCES

[1] S.Ullman and R.Basri ; " Recognition by Linear Combination of Models", IEEE Trans. Patt. Anal. Machine Intell., Vol.13, No.10, pp.992-1006, 1991.

[2] S.J.Dickinson, et al. ; "3D Shape Recovery Using Distributed Aspect Matching", IEEE Trans. Patt. Anal. Machine Intell., Vol.14, No.2, pp.174-198, 1992.

[3] M. Dhome, et al. ;"Determination of the Attitude of 3D Objects from a Single Perspective View," IEEE Trans. Patt. Anal. Machine Intell., Vol.11, No.12, pp.1265-1278, 1989.

[4] D.G.Lowe ; " Three-Dimensional Object Recognition from Single Two-Dimensional Images," Artificial Intell., Vol.31, pp.355-395, 1987.

[5] D.G.Lowe ; "Fitting Parameterized Three-Dimensional Models to Images," IEEE Trans. Patt. Anal. Machine Intell., Vol.13, No.5, pp.441-450, 1991.

[6] K.Ikeuchi and T.Kanade ; "Automatic Generation of Object Recognition Programs," Proc. IEEE, Vol.76, No.8, pp.1016-1035, 1988.